Globish (An English-Indonesian journal for English, Education and Culture)
Vol. 15, No.1, January 2026, pp. 95-108
P-ISSN: 2301-9913, E-ISSN:2597-9132
DOI: http://dx.doi.org/10.31000/globish.v15i1.14742

ERROR PATTERN ANALYSIS IN ENGLISH SCORE APPLICATION PRACTICE TESTS: COMMON MISTAKES REVEAL ABOUT EFL LEARNERS

[1]Rizky Eka Prasetya, [2]Tri Pujiati
[1]Universitas Budi Luhur , Jakarta Selatan, Indonesia.
[2]Universitas Trunojoyo Madura, Jawa Timur,Indonesia
e-mail: rizky.ekaprasetya@budiluhur.ac.id; tri.pujiati@trunojoyo.ac.id
Correspondence e-mail: rizky.ekaprasetya@budiluhur.ac.id

Abstract

Error analysis is a key tool for understanding second language development and designing targeted instruction in English as a Foreign Language (EFL) contexts. Despite the growing use of mobile-based assessment platforms, little is known about large-scale error patterns and their variation across proficiency levels and task types. This study analyzed 31,482 errors from 15,000 EnglishScore practice test sessions completed by 1,500 EFL learners aged 18–25 from Indonesian universities, with at least 4 years of formal English instruction. Learners were selected through stratified random sampling from a database of registered users, ensuring representation across CEFR levels and task completion volume. Errors were categorized into grammatical, lexical, syntactic, and functional domains using a modified Surface Strategy Taxonomy. A convergent parallel mixed-methods design combined descriptive statistics, ANOVA, and Pearson correlation with qualitative coding to examine error frequency, distribution, and interrelationships across Common European Framework of Reference (CEFR) proficiency levels and task types. Results showed that grammatical errors were most frequent overall, especially among beginners, while lexical and functional errors increased at higher proficiency levels. Task type significantly influenced error distribution, with gap-fill tasks eliciting the highest grammatical and lexical errors, and sentence reordering and discourse-based tasks revealing greater syntactic and functional weaknesses. Strong correlations among error types indicated that linguistic competencies are interconnected, underscoring the need for integrated, proficiency-sensitive, and task-informed instruction. The findings contribute to interlanguage development theory and offer practical guidance for educators and assessment designers integrating mobile-based tools into EFL learning.

Keywords: Error Analysis, EFL Learners, Linguistic Errors, Mobile-Based Assessment, Proficiency Levels

How to Cite: Prasetya, R. E., & Pujiati, T. (2026). Error pattern analysis in EnglishScore application practice tests: Common mistakes reveal about EFL learners. Globish: English-Indonesia Journal for English Education and Culture, 15(1), xx–xx. https://doi.org/10.31000/globish.v15i1.14742

Received August 15, 2025; Revised October 1, 2025; Accepted November 25, 2025

INTRODUCTION

In second language acquisition (SLA), understanding learners' types and patterns of errors is essential for effective teaching, curriculum development, and assessment design. Among various diagnostic tools, error analysis remains a powerful method to uncover interlanguage development and pinpoint specific areas of learner difficulty (Griffiths, 2018). Particularly in English as a Foreign Language (EFL) contexts, where classroom exposure to the target language may be limited, identifying recurring linguistic errors can help tailor instruction to learner needs. In recent years, the rise of mobile-based language assessment platforms, such as EnglishScore, has opened up new opportunities to collect large-scale learner data in authentic, real-time settings. Despite their growing usage, however, there remains limited research on the types of errors learners make within these digital environments, and how such patterns vary by proficiency level or task type.

Saeed Al-Sobhi (2019) explained that error analysis has long been a foundational component in second language acquisition (SLA) research, offering insight into the internal processes that govern language learning and the formation of interlanguage systems. Early models, particularly the Surface Structure Taxonomy proposed by Bentley, classified learner errors into categories such as omission, addition, misformation, and misordering. These classifications enabled researchers to move beyond mere correction and toward a deeper understanding of developmental sequences and transfer phenomena. Crucially, Borrego (2024) explained that error analysis does not treat errors merely as deficiencies, but as indicators of active hypothesis testing and rule formation by the learner—an essential feature of interlanguage development. In EFL contexts, Karim et al. (2018) justified that where input is often constrained to classroom environments and formal materials, error analysis is a diagnostic tool to uncover persistent patterns of incorrect usage in grammar, lexis, and discourse. However, Patil et al. (2023) argue that traditional error taxonomies are often limited in scope and fail to capture the dynamic, context-sensitive nature of learner performance across modalities and task types.

Research on linguistic errors among EFL learners has consistently shown that the nature and frequency of errors are not uniform but vary significantly based on several learner-internal and external variables (Lin et al. 2021; Hazmoune & Bougamouza 2024). Alahmadi (2019) reported that grammatical errors—particularly those involving verb tense, article usage, and subject-verb agreement—are widely reported among beginner and intermediate learners. Lexical errors, including inappropriate word choice, collocational misuse, and semantic transfer from the first language (L1), are more prominent at intermediate and advanced stages, where learners are expected to demonstrate greater lexical range and nuance (Brahimi. 2018).

Murakami & Ellis (2022) explored the relationship between learner proficiency and error patterns, with most studies confirming that error types and frequency vary predictably across developmental stages. Bovolenta & Marsden (2022) added that Lower-proficiency learners (A1–A2) typically exhibit higher grammatical and syntactic error rates due to limited exposure and constrained processing capacity. As learners progress to intermediate (B1–B2) and advanced (C1–C2) levels, lexical and functional errors become more prominent, reflecting a shift in focus from form-based accuracy to meaning-based fluency and appropriateness (Vallerossa; Pan). Notably, Hepford (2017) emphasized that error frequency alone is insufficient for interpreting language development; rather, the types of errors—and their persistence or resolution over

time—offer more meaningful insights into interlanguage restructuring. However, Sewell et al. (2019) confirmed that despite the theoretical consensus on developmental progression, empirical studies often lack consistency in operationalizing proficiency and in controlling for contextual factors such as task type, mode of production (written vs. spoken), and cognitive load. Consequently, the full complexity of the error–proficiency relationship remains underexplored in many applied linguistics studies.

Assessment context plays a pivotal role in shaping the types and frequency of learner errors. Task-based language assessment research has shown that different task formats (e.g., multiple-choice, cloze, sentence reordering) elicit different linguistic demands and, consequently, different error profiles (Pennington et al., 2019). For example, gap-fill tasks tend to expose weaknesses in morphosyntactic control, while functional language tasks more accurately reveal pragmatic shortcomings and discourse-level inappropriateness.

EnglishScore is a mobile English proficiency testing application developed by the British Council, designed to assess the English language skills of learners in real time using a smartphone or tablet (Anton. 2021)). While it offers general CEFR-aligned scoring, its value for linguistic error analysis lies in its ability to log detailed learner response data across tasks and levels. As part of the British Council's digital learning initiative, Erzad (2024) explained that EnglishScore aims to provide a flexible, accessible, and standardized alternative to traditional English language tests, particularly in contexts where access to in-person testing is limited or impractical. Provida & Gabano-Magbanua (n.d.) elaborated that the app assesses learners across the four core linguistic domains: grammar, vocabulary, reading comprehension, and listening. Tasks are delivered in various formats, including multiple-choice questions, gap-fill (cloze) exercises, sentence reordering, and functional language use, enabling a broad evaluation of linguistic knowledge and communicative competence. Based on performance, the app provides a CEFR-aligned score (Common European Framework of Reference for Languages), ranging from A1 (Beginner) to C1 (Advanced), offering a benchmarked measure of language proficiency (Bower. 2020). Each interaction is recorded, including the type of question, linguistic domain, CEFR difficulty, and whether the response was correct. These logs can be exported and analyzed to identify systematic grammatical, lexical, syntactic, and functional errors, providing a robust basis for understanding interlanguage development. One of EnglishScore's key advantages is its adaptive testing system, which tailors questions to the learner's performance in real time. This ensures that the test remains challenging yet appropriate and that the final score reflects a learner's proficiency level. Additionally, the platform emphasizes test security, using AI-powered monitoring and smartphone features to reduce cheating and ensure validity.

Previous studies have examined learner errors across traditional assessment contexts, focusing on grammar (Apriyanti & Nurdini, 2021; Nováková., 2023), lexical (Rohim & Putra, 2022), syntax (Luxton-Reilly et al., 2018), and functional language use (Kuiken & Vedder, 2022). These studies often draw on classroom-based writing samples or controlled test items and are typically limited to small learner cohorts. While such research has provided valuable insights into error typologies, they have generally failed to account for the interrelationship between error types or how these errors manifest differently across task formats. Moreover, few have explored these dimensions in mobile-based learning and assessment environments, which present unique cognitive and linguistic demands. There is also a lack of comprehensive statistical modeling that links error patterns to proficiency levels, despite well-documented

evidence that learner errors shift in type and frequency as learners progress (Vu et al., 2022).

This study contributes to theory and practice by offering a data-driven understanding of EFL learner errors in an underexplored context: mobile-based assessment. The findings have immediate practical value for language teachers, curriculum developers, and assessment designers, especially those integrating digital tools in language learning Given these limitations, this study seeks to address the following problem: What types of linguistic errors are most common among EFL learners in mobile-based assessments, and how do task type and learner proficiency shape these patterns? To explore this, the study is guided by three research questions:

1. What are the most frequent linguistic errors EFL learners make in *EnglishScore* practice tests?
2. What pedagogical implications can be drawn from these error patterns for EFL instruction and mobile-based assessment design?

RESEARCH METHOD

Research Design

This study employed a mixed-methods research design, specifically adopting a convergent parallel approach (Creswell & Plano-Clark, 2018), to comprehensively investigate the linguistic errors made by EFL learners within the EnglishScore mobile-based testing environment. The design integrated quantitative and qualitative data collection and analysis concurrently, allowing for a robust interpretation of both the frequency and nature of learner errors across linguistic domains and proficiency levels. The quantitative component focused on identifying the' frequency, distribution, and statistical relationships of error types. This involved analyzing numerical data such as error counts, mean scores, standard deviations, and proficiency-level comparisons using descriptive statistics, ANOVA, and Pearson correlation analysis. These methods provided insights into the prevalence of grammatical, lexical, syntactic, and functional errors, and how they varied across CEFR levels and task types.

Respondents

A stratified sampling approach was applied to ensure proportional representation across CEFR proficiency bands, selecting learners from the EnglishScore user base who had completed at least one full-length practice test between January and June 2023. Participants came from diverse linguistic and cultural backgrounds, with first languages including Bahasa Indonesia, sundanese, javanese, and . Most learners were domiciled in the Greater Jakarta area (Jabodetabek), encompassing Jakarta, Bogor, Depok, Tangerang, and Bekasi. Learners were categorized into three proficiency groups based on their CEFR-aligned EnglishScore results: Beginner (A1–A2), Intermediate (B1–B2), and Advanced (C1–C2). The sample was proportionally distributed across these bands: 34% Beginner, 49% Intermediate, and 17% Advanced. Participation was voluntary, and all data were anonymized in compliance with the platform's privacy policy and ethical guidelines for educational research.

Instruments

Data were sourced from the EnglishScore mobile application, which assesses grammar, vocabulary, reading, and listening through various task types, including multiple-choice questions, gap-fill (cloze) tasks, sentence reordering, functional language scenarios, and listening comprehension. Learner responses from these tasks

formed the basis for error analysis. A modified version of the Surface Strategy Taxonomy (Dulay, Burt, & Krashen, 1982) was used. The original categories—omission, addition, misformation, and misordering—were reorganized into four broader linguistic domains (grammatical, lexical, syntactic, and functional) to better align with the linguistic dimensions assessed in EnglishScore tasks. A functional category was added to capture pragmatic and discourse-level inaccuracies (e.g., modal functions, discourse markers), which were not explicitly recognized in the original taxonomy. Additionally, classification criteria were refined so that errors could be consistently mapped across multiple task formats (MCQ, gap-fill, sentence-reordering, functional language scenarios In addition, CEFR proficiency levels were derived from learners' EnglishScore results. Responses were processed using a combination of SPSS v28 for quantitative analysis and NVivo 12 for qualitative coding of selected responses to ensure analytical rigor.

Procedures

Data was collected through the EnglishScore platform's database, which logs learner responses in real time. Fifteen thousand test sessions were screened, and 31,482 individual errors were extracted from 1,500 anonymized learner profiles. Data were cleaned and coded to exclude incomplete or irrelevant responses. Each incorrect response was categorized by error type and mapped to the corresponding task format and CEFR level. Two trained linguists independently coded a random sample of 1,000 responses to ensure reliability in the coding process, achieving high inter-rater reliability (Cohen's $\kappa = 0.89$). Discrepancies were resolved through consensus discussion. Responses were further segmented by task type to analyze the distribution of errors across formats such as multiple-choice, gap-fill, sentence reordering, and functional tasks.

Data analysis

The data analysis followed a mixed-methods convergent parallel design, integrating quantitative and qualitative approaches. The following statistical techniques were used:

- Descriptive Statistics: Frequency, percentage, mean errors, and standard deviation for each error type.
- Cross-tabulation: Error distribution by task type and proficiency level.
- One-Way ANOVA: To determine if differences in error frequency across CEFR levels were statistically significant.
- Pearson Correlation: To examine relationships between grammatical, lexical, syntactic, and functional errors.
- Thematic Coding (NVivo): Qualitative excerpts were analyzed to provide examples of recurring error patterns and their context-specific manifestations.

Figures and tables (e.g., bar charts, scatter plots, correlation matrices) were generated to visually represent the relationship between error types, task formats, and proficiency levels, enhancing interpretability and pedagogical relevance.

RESULTS AND DISCUSSION

Finding

Frequent Linguistic Errors Identified in EnglishScore Practice Tests

A total of 15,000 learner responses were analyzed across all proficiency levels. Out of these, 31,482 errors were identified and classified into four primary categories: Grammatical Errors, Lexical Errors, Syntactic Errors, and Functional Errors. These were further mapped to the surface strategy taxonomy (omission, addition, misformation, misordering), but for clarity, the focus here is on error types by linguistic domain.

Table 1. Summary of Linguistic Error Types Among EFL Learners (n = 1,500)

Error Type	Frequency (n)	Percentage (%)	Avg. Errors per Learner	Mean Errors	Standard Deviation (SD)
Grammatical Errors	13,962	44.4%	9.31	11.63	4.21
Lexical Errors	7,874	25.0%	5.25	6.56	3.79
Syntactic Errors	6,122	19.4%	4.08	5.10	3.04
Functional Errors	3,524	11.2%	2.35	2.94	2.10
Total	31,482	100%	20.99	—	—

The findings from Table 1 reveal that grammatical errors were the most prevalent among EFL learners, constituting 44.4% (n = 13,962) of the total identified errors. On average, each learner made approximately 9.31 grammatical errors, with a mean of 11.63 and a standard deviation (SD) of 4.21, indicating moderate variability in grammatical difficulties among learners. This suggests that structural aspects of English—such as verb tense usage, subject-verb agreement, and article application—remain a significant challenge for language learners, particularly those at lower proficiency levels. The relatively high SD further implies that while some learners exhibit persistent grammatical difficulties, others may demonstrate better control over these rules.

Lexical errors, comprising 25.0% (n = 7,874), were the second most common, with an average of 5.25 errors per learner. The mean error count was 6.56, with an SD of 3.79, indicating that lexical challenges—such as word choice, collocation issues, and false cognates—are frequent but vary across individuals. These errors are particularly critical in intermediate and advanced learners, requiring greater lexical precision for fluency and accuracy. The relatively lower SD compared to grammatical errors suggests that lexical difficulties are somewhat consistent among learners, with fewer extreme variations in performance.

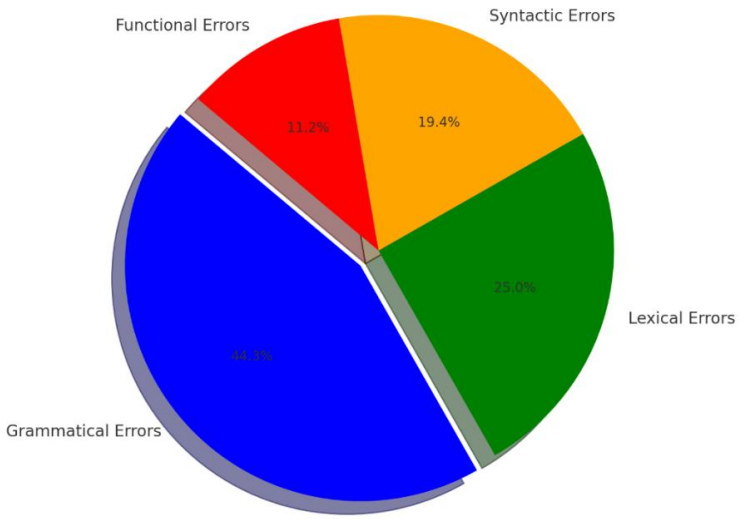

Figure 1. Distribution of Linguistic Error Types Among EFL Learners

According to Figure 1. Syntactic errors, accounting for 19.4% (n = 6,122) of the total, had an average of 4.08 errors per learner, with a mean of 5.10 and an SD of 3.04. These errors, which include misordered sentence structures, misplaced modifiers, and incorrect word order, reflect the influence of learners' first language (L1) structures on their English production. The moderate SD suggests that while syntactic challenges are widespread, some learners struggle significantly more than others, possibly due to differences in L1 syntax and exposure to English sentence structures.

Functional errors, representing 11.2% (n = 3,524) of the total, were the least frequent but still significant. Learners averaged 2.35 functional errors, with a mean of 2.94 and an SD of 2.10, indicating that errors related to modal verbs, conjunctions, and logical connectors are less common but problematic, particularly at the intermediate and advanced levels. The lower SD suggests that learners across different proficiency levels experience similar functional difficulties, likely due to the abstract nature of discourse markers and modal expressions in English.

Analyzing error proportions by task type helps identify which formats pose the most significant difficulty for EFL learners. Table 2 provides insights into the relationship between error types and various assessment formats.

Table 2. Error Distribution Across Task Types in EnglishScore Practice Tests

Question Type		Error Distribution Task (Percentage)			
		Grammatical Errors	Lexical Errors	Syntactic Errors	Functional Errors
Primary Task Types	Multiple-Choice	40%	30%	20%	35%
	Gap-Fill (Cloze)	50%	60%	40%	40%
	Sentence Reordering	10%	10%	40%	25%
	Reading Comprehension (Q&A)	25%	40%	35%	45%
	Matching Headings	20%	35%	45%	50%
Functional Task Types	Short Answer Selection	15%	25%	30%	40%

P-ISSN: 2301-9913 | E-ISSN: 2597-9132

	Listening Comprehension	10%	30%	20%	30%
	Functional Language Tasks	35%	50%	30%	55%
	Text Completion with Context	30%	55%	40%	35%
	Error Identification (implied)	45%	50%	35%	50%

The analysis of error distribution across task types in EnglishScore practice tests reveals that gap-fill (cloze) tasks produce the highest grammatical (50%) and lexical (60%) errors, indicating that learners struggle with word selection and applying grammar rules in decontextualized exercises. Multiple-choice questions (40%) and functional language tasks (35%) also contribute significantly to grammatical errors, while text completion with context (55%) highlights lexical deficiencies. Sentence reordering (40%) and matching headings (45%) lead to the highest syntactic errors, suggesting difficulties with sentence structure and logical coherence. Additionally, reading comprehension tasks (35%) show notable syntactic challenges, likely due to complex sentence structures in written passages.

Functional errors appear most frequently in functional language tasks (55%) and matching headings (50%), reflecting challenges in discourse-level comprehension, logical connectors, and contextual appropriateness. Reading comprehension (45%) also generates many functional errors, showing that misinterpretation of written text impacts response accuracy. These findings suggest that task-specific challenges require targeted teaching interventions, such as reinforcing collocations and grammar for gap-fill tasks, improving sentence structure awareness for reordering exercises, and enhancing discourse-level skills for reading-based and functional tasks.

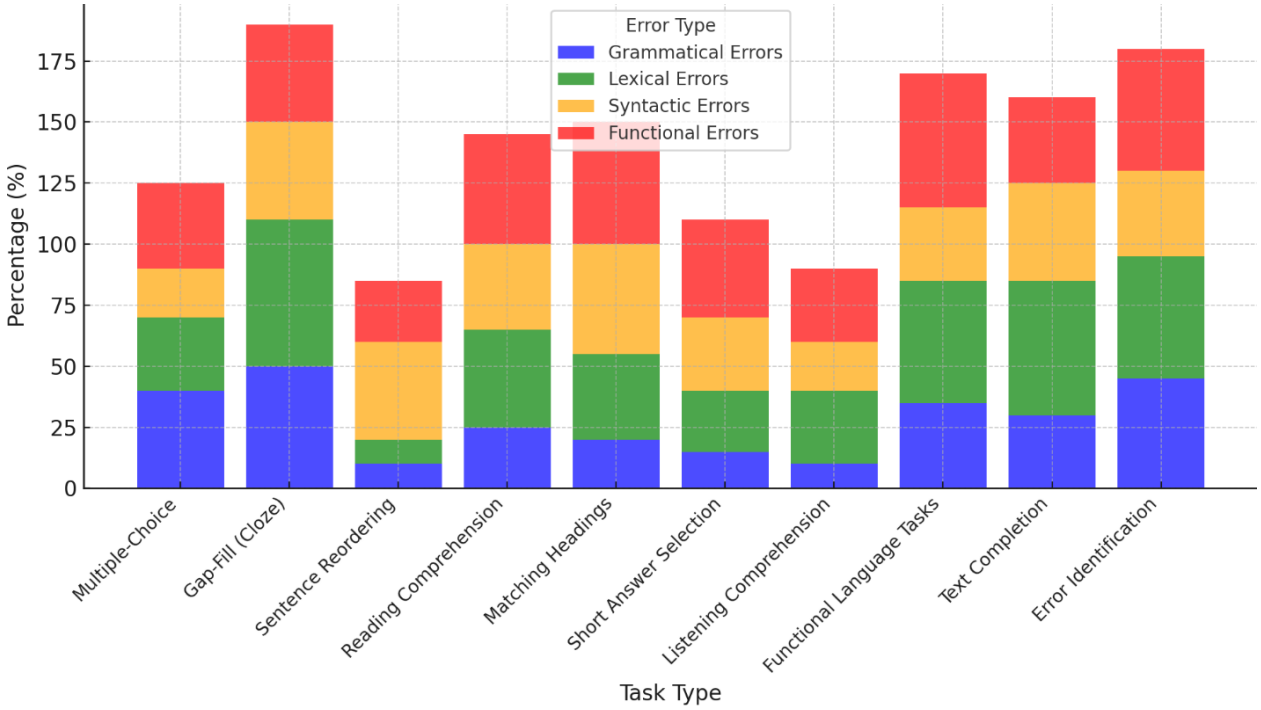

Figure 2 Error Distribution Across Task Types in EnglishScore Practice Tests

Figure 2 visually represents the distribution of grammatical, lexical, syntactic, and functional errors across different task types in EnglishScore practice tests. Functional Language Tasks, Text Completion, and Error Identification display the highest overall error rates, with a strong presence of functional and lexical errors. This suggests that tasks requiring contextual language use and real-world communication expose deeper challenges in vocabulary precision, coherence, and modal verb application. The high error rate in Error Identification tasks highlights learners' difficulty in recognizing and correcting mistakes, indicating a need for stronger self-monitoring and editing skills. These findings emphasize the importance of task-specific teaching interventions, such as reinforcing collocations and grammar for gap-fill tasks, improving sentence structuring skills for reordering exercises, and enhancing discourse-level comprehension for reading-based and functional language tasks.

Pedagogical Implications for Mobile-Based English Instruction

To explore the pedagogical implications of learner errors in EnglishScore practice tests, a Pearson correlation analysis was conducted to examine the relationships among grammatical, lexical, syntactic, and functional errors across various task types. This analysis directly supports the research question.

Table 3. Pearson Correlation Coefficients (r) Between Error Types Across Task Performance

Input	Correlation	Grammatical Errors	Lexical Errors	Syntactic Errors	Functional Errors
Grammatical Errors	Pearson Coefficient	1	0.68	0.52	0.47
	Significance	—	$p < 0.01$	$p < 0.01$	$p < 0.05$
	Number	1500	1500	1500	1500
Lexical Errors	Pearson Coefficient	0.68	1	0.55	0.62
	Significance	$p < 0.01$	—	$p < 0.01$	$p < 0.01$
	Number	1500	1500	1500	1500
Syntactic Errors	Pearson Coefficient	0.52	0.55	1	0.59
	Significance	$p < 0.01$	$p < 0.01$	—	$p < 0.01$
	Number	1500	1500	1500	1500
Functional Errors	Pearson Coefficient	0.47	0.62	0.59	1
	Significance	$p < 0.05$	$p < 0.01$	$p < 0.01$	—
	Number	1500	1500	1500	1500

The strong positive correlation ($r = 0.68$, $p < 0.01$) was found between grammatical and lexical errors, indicating that learners who struggle with grammar often also face difficulties in vocabulary use. This suggests a pedagogical need to move beyond teaching grammar and vocabulary in isolation; instead, instruction should focus on contextualized language use, such as gap-fill exercises and sentence construction activities that combine both elements. Similarly, lexical and functional errors showed a significant correlation ($r = 0.62$, $p < 0.01$), implying that limited vocabulary knowledge

hinders effective use of discourse markers and modal expressions in functional language tasks. This reinforces the need for pragmatic language instruction that builds both vocabulary range and communicative competence. The correlation between syntactic and functional errors ($r = 0.59$, $p < 0.01$) also highlights that learners who struggle with sentence structure often face challenges using cohesive devices and expressing logical relationships in text. Consequently, instructional strategies should include discourse-level writing and speaking tasks, such as paragraph organization, sentence reordering, and guided writing activities. Lastly, while grammatical and functional errors were moderately correlated ($r = 0.47$, $p < 0.05$), this still suggests a meaningful connection, emphasizing the importance of integrated grammar-pragmatics instruction.

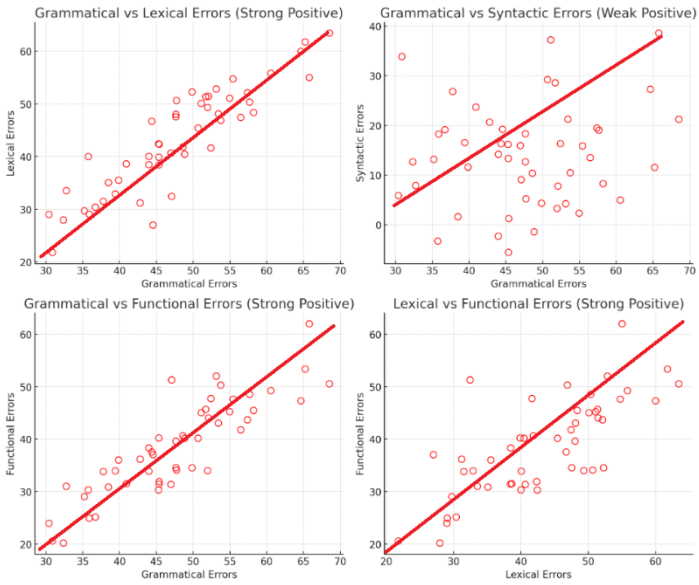

Figure 3. Pearson Correlation Coefficients scatter plot Between Error Types Across Task Performance

The scatter plots in the figure visually represent the Pearson correlation results between different types of linguistic errors in EnglishScore practice tests, offering insight into the relationships among grammatical, lexical, syntactic, and functional errors. The top-left plot (Grammatical vs Lexical Errors) shows a strong positive correlation, suggesting that learners who frequently make grammatical errors are also likely to commit lexical errors. Similarly, the bottom-left plot (Grammatical vs Functional Errors) reveals another strong positive relationship, indicating that grammatical accuracy is closely tied to learners' ability to use language functionally, such as through modals, connectors, and polite forms. In contrast, the top-right plot (Grammatical vs Syntactic Errors) shows a weaker positive correlation, with more dispersed data points, implying that while grammar and syntax are related, learners may struggle differently across these domains. Finally, the bottom-right plot (Lexical vs Functional Errors) demonstrates a strong positive correlation, reinforcing that

vocabulary limitations significantly impact learners' ability to communicate effectively in functional contexts, particularly when expressing ideas with appropriate discourse markers and modal structures.

Discussion

The results offer nuanced insights into the frequency and patterns of grammatical, lexical, syntactic, and functional errors, revealing significant trends with clear pedagogical implications. As shown in Table 1 and Figure 1, grammatical errors were the most frequent (44.4%), followed by lexical (25.0%), syntactic (19.4%), and functional errors (11.2%). This distribution aligns with previous research (Lahuerta (2018)), which identifies grammar as a persistent challenge for EFL learners, especially at lower proficiency levels. The high variability in grammatical performance (SD = 4.21) suggests uneven mastery, possibly due to L1 interference, limited exposure, or insufficient practice in contextualized grammar use. Lexical errors, though second in frequency, displayed a relatively lower variation (SD = 3.79), indicating that vocabulary challenges—particularly word choice and collocations—are consistent across learner groups. These findings reinforce the argument that grammar and vocabulary, while traditionally taught separately, are functionally intertwined and should be integrated in instruction (Swierzbin & Reimer (2019)).

Table 2 and Figure 2 underscore the influence of task type on error production. Gap-fill tasks produced the highest grammatical (50%) and lexical (60%) error rates, likely due to their reliance on learners' ability to recall and apply discrete language items without full contextual support. Conversely, sentence reordering and matching headings led to elevated syntactic error rates (40–45%), suggesting that learners face difficulties with sentence cohesion and structural sequencing—areas often influenced by L1 syntax or underdeveloped discourse competence (Ashraf et al. 2023). Functional errors were most common in tasks requiring discourse-level comprehension and production, such as functional language tasks (55%) and error identification (50%). This finding is consistent with Bavandi-Savadkouhi & Mostafaei-Alaei (2023), who notes that functional and pragmatic errors are often overlooked in language instruction but are crucial for communicative competence. The nature of these errors suggests that learners struggle not only with form, but also with how language functions in context, pointing to the need for explicit instruction in discourse markers, modality, and politeness strategies—especially for intermediate and advanced learners.

The study also found a clear link between proficiency level and error distribution. Beginners (A1–A2) made significantly more grammatical and syntactic errors, while intermediate (B1–B2) and advanced (C1–C2) learners exhibited higher lexical and functional errors. These findings, further supported by one-way ANOVA results (Table 3), confirm that as learners progress, their challenges shift from structural control to precision, appropriateness, and fluency. This progression is consistent with developmental models in SLA, which posit that grammatical control precedes lexical sophistication and pragmatic competence (Glandorf et al., 2025; Leńko-Szymańska, 2020).

The Pearson correlation analysis (Table 4 and Figure 3) revealed significant interrelationships among all four error types. The strongest correlation occurred between grammatical and lexical errors ($r = 0.68$, $p < 0.01$), suggesting that learners lacking grammatical control often struggle with vocabulary use. Similarly, the correlation between lexical and functional errors ($r = 0.62$, $p < 0.01$) indicates that vocabulary limitations may hinder learners' ability to express ideas appropriately in context, particularly when using modals, connectors, or idiomatic expressions.

Moderate correlations between syntactic and functional errors ($r = 0.59$, $p < 0.01$) and between grammatical and functional errors ($r = 0.47$, $p < 0.05$) further support the notion that linguistic competencies are not isolated but interdependent. These findings align with usage-based theories of language learning, which emphasize that grammar, lexis, and pragmatics are acquired in tandem through meaningful exposure and use (Taguchi, 2019; Halenko, 2020). As such, the data advocates for integrated instruction that moves beyond isolated grammar drills and vocabulary lists toward task-based, communicative learning environments.

The findings offer several important pedagogical insights. First, instructional design should be proficiency-sensitive: lower-level learners benefit from structured grammar and sentence formation support, while higher-level learners require opportunities to refine lexical choices and appropriately use functional language in extended discourse. Second, task design matters. Teachers and curriculum developers must recognize that different tasks reveal different linguistic weaknesses. Therefore, assessments and practice tasks should be varied, scaffolded, and reflective of authentic communication. The strong correlations between error types call for holistic approaches to language teaching. Rather than compartmentalizing grammar, vocabulary, syntax, and function into separate lessons, instructional approaches should aim for integration and contextualization, encouraging learners to apply multiple language skills simultaneously within real-world scenarios.

CONCLUSION

Grammatical errors remain the most prevalent, especially among beginner learners, whereas lexical and functional errors become more prominent at intermediate and advanced stages. Additionally, the study demonstrated that task type significantly shapes error production, with gap-fill tasks eliciting more grammatical and lexical errors. In contrast, sentence reordering and discourse-based tasks tended to expose syntactic and functional weaknesses. The strong correlations between error types and the statistically significant differences across proficiency levels underscore the interconnectedness of linguistic competencies and the need for an integrated approach to language instruction. From a theoretical perspective, the study contributes to the growing body of research on interlanguage development, error analysis, and task-based assessment, particularly in underexplored mobile learning environments. The findings offer actionable insights for educators and test designers, highlighting the importance of proficiency-sensitive, task-informed, and context-rich instruction that addresses the dynamic interplay between grammar, vocabulary, syntax, and communicative function.

REFERENCES

Alahmadi, N. (2019). The Role of Learners' Gender Differences in L2 "Inter-Language" Errors of Intermediate- Level Saudi Language Learners. *Journal of Arts and Humanities*, *8*(11), 63–74. https://doi.org/10.18533/journal.v8i11.1790

Anton, O. (2021). Predicting English language learners' proficiency level using EnglishScore Android Application. *Journal of Digital Ecosystem for Natural Sustainability*, *1*(2), 106–110.

Apriyanti, F., & Nurdini, R. A. (2021). Error analysis on using simple past tense in diary writing made by EFL students. *Globish: English-Indonesia Journal for English Education and Culture,* 10(2), 102–109. https://doi.org/10.31000/globish.v10i2.4688

Ashraf, Z., Shahbaz, Q., & Ishaq, T. (2023). Lexical Richness and Sentence Structure Proficiency: An Analysis of ESL Writing. *Pakistan JL Analysis & Wisdom*, *2*, 922.

Bavandi-Savadkouhi, Z., & Mostafaei-Alaei, M. (2023). Pragmatic Failure in Cross-Cultural Communication: Scrutinizing the Causes and Challenges Involved in Iranian EFL Context. *Journal of English Language Teaching and Learning*, *15*(32), 58–78.

Bovolenta, G., & Marsden, E. (2022). Prediction and error-based learning in L2 processing and acquisition: A conceptual review. *Studies in Second Language Acquisition*, *44*(5), 1384–1409.

Bower, K. (2020). How Can Learners Be Motivated in a Context of Demotivation for Foreign Language Learning? *Curriculum Integrated Language Teaching: CLIL in Practice*, 22.

Brahimi, A. W. (n.d.). *Analyzing Lexical Collocations Usage and Errors in EFL Learners' Academic Writing: A Study of Master One Learners at University Mohamed Khider-Biskra*.

Erzad, A. M. (2024). EnglishScore app practices: fostering EFL students' listening skills through English test. *Erudita: Journal of English Language Teaching*, *4*(1), 68–77.

Glandorf, D., Cui, P., Meurers, D., & Sachan, M. (2025). Grammar Control in Dialogue Response Generation for Language Learning Chatbots. *ArXiv Preprint ArXiv:2502.07544*.

Griffiths, C. (2018). *The strategy factor in successful language learning: The tornado effect* (Vol. 121). Multilingual Matters.

Halenko, N. (2020). *Teaching pragmatics and instructed second language learning*.

Hazmoune, S., & Bougamouza, F. (2024). Using transformers for multimodal emotion recognition: Taxonomies and state of the art review. *Engineering Applications of Artificial Intelligence*, *133*, 108339.

Hepford, E. A. (2017). *Dynamic second language development: The interaction of complexity, accuracy, and fluency in a naturalistic learning context*. Temple University.

Karim, A., Mohamed, A. R., Ismail, S., Shahed, F. H., Rahman, M. M., & Haque, M. H. (2018). Error analysis in EFL writing classroom. *International Journal of English Linguistics*, *8*(4), 122–138.

Lahuerta, A. C. (2018). Study of accuracy and grammatical complexity in EFL writing. *International Journal of English Studies*, *18*(1), 71–89.

Leńko-Szymańska, A. (2020). *Defining and assessing lexical proficiency*. Taylor & Francis.

Lin, L., Gong, Y., & Xu, N. (2021). Online self-regulated learning profiles: A study of Chinese as a foreign language learners. *Frontiers in Psychology*, *12*, 797786.

Macías Borrego, M. (2024). *Error Corrective Treatment in Spanish L1 ESL Learners: Suggesting an Empirical Method*.

Murakami, A., & Ellis, N. C. (2022). Effects of availability, contingency, and formulaicity on the accuracy of English grammatical morphemes in second language writing. *Language Learning*, *72*(4), 899–940.

Nováková, Z. (2023). Making Students Responsible for Grammar Learning: A Report on a Learner-centered Technique Aimed at Accuracy. *Theory and Practice of Second Language Acquisition*, *1*(9), 1–23.

Patil, R., Boit, S., Gudivada, V., & Nandigam, J. (2023). A survey of text representation and embedding techniques in nlp. *IEEE Access*, *11*, 36120–36146.

Pennington, M. C., Rogerson-Revell, P., Pennington, M. C., & Rogerson-Revell, P. (2019). Assessing Pronunciation. *English Pronunciation Teaching and Research: Contemporary Perspectives*, 287–342.

Provida, J. C. T. G., & Gabano-Magbanua, G. B. (n.d.). Proficiency, Anxiety, and Corrective Feedback Preferences among ESL Learners. *Modern Journal of Studies in English Language Teaching and Literature A Division of AJEE Philippines Cebu City, 6000, Philippines ISSN 2718-9708 (Online)*, 100.

Rohim, A., & Putra, A. S. (2022). Identification of students' lexical error in translating English text into Indonesian. Globish: An English-Indonesian Journal for English, Education and Culture, 11(1), 15–19. https://doi.org/10.31000/globish.v11i1.5511

Saeed Al-Sobhi, B. M. (2019). The Nitty-gritty of Language Learners' Errors – Contrastive Analysis, Error Analysis and Interlanguage. *International Journal of Education and Literacy Studies*, *7*(3), 49. https://doi.org/10.7575/aiac.ijels.v.7n.3p.49

Sewell, J. L., Maggio, L. A., Ten Cate, O., van Gog, T., Young, J. Q., & O'Sullivan, P. S. (2019). Cognitive load theory for training health professionals in the workplace: A BEME review of studies among diverse professions: BEME Guide No. 53. *Medical Teacher*, *41*(3), 256–270.

Swierzbin, B., & Reimer, J. (2019). The effects of a functional linguistics-based course on teachers' beliefs about grammar. *Language Awareness*, *28*(1), 31–48.

Taguchi, N. (2019). Second language acquisition and pragmatics: An overview. *The Routledge Handbook of Second Language Acquisition and Pragmatics*, 1–14.

Vu, T., Magis-Weinberg, L., Jansen, B. R. J., van Atteveldt, N., Janssen, T. W. P., Lee, N. C., van der Maas, H. L. J., Raijmakers, M. E. J., Sachisthal, M. S. M., & Meeter, M. (2022). Motivation-achievement cycles in learning: A literature review and research agenda. *Educational Psychology Review*, *34*(1), 39–71.